PATENT SPECIFICATION (11) 1 491 702

1 491 702

(21) Application No. 18987/74 (22) Filed 1 May 1974
(23) Complete Specification filed 1 April 1975
(44) Complete Specification published 16 Nov. 1977
(51) INT. CL.² G11C 29/00
(52) Index at acceptance
G4A 12C 12N 13E 6M1 6MX 6X EF
(72) Inventors JOHN MARTIN HARPER
IVOR EDWIN REYNOLDS
NIGEL RONALD HASSALL BAILEY

(19)

(54) IMPROVEMENTS IN OR RELATING TO MEMORY SYSTEMS

(71) We, INTERNATIONAL COMPUTERS LIMITED, a British Company, of ICL House, Putney, London, S.W.15, do hereby declare the invention for which we pray that a patent may be granted to us, and the method by which it is to be performed, to be particularly described in and by the following statement:—

This invention relates to memory systems. Several different solutions have been proposed for the problem which is created by faults in memory devices. One class of solutions assumes that some degree of data corruption will occur and codes the data in a redundant system, such as a Hamming code, which will provide detection of data errors and also reconstruction of the correct data within limits. Whilst such arrangements are effective and are widely used, they require a substantial amount of extra equipment and also additional time whenever data has to be reconstructed.

A second class of solutions attemps to to avoid the problem by testing a memory plane after manufacture to determine which memory positions are faulty. The wiring of the memory plane and/or of the address selection device is then carried out in such a way that the faulty positions are not used. The main disadvantage is that this solution cannot deal with faults which arise after the wiring has been effected. There is also difficulty in devising procedures which ensure that all possible combinations of conditions which may produce a fault are tested.

A third class of solutions provides a limited amount of additional memory capacity which is substituted as required for faulty positions. This substitution may be made initially for faulty positions which are discovered during testing, or it may be made on a continuing basis as faults are discovered. This latter basis is clearly the most flexible way of dealing with faults. The invention seeks to provide an improved arrangement for implementing such a facility.

According to the invention a memory system comprises: a main memory having a plurality of addressable locations each arranged to store an item of data, some of which locations may be faulty; an address register for addressing the main memory to select a storage location; a data checking device responsive to a data item read out from a selected storage location to produce an error output if the data item read out contains an error indicating that the location is faulty; an auxiliary memory having fewer addressable locations than the main memory, each arranged to store an item of data; address code conversion means coupled to the address register and responsive to said error output arranged to assign an address of a replacement location in the auxiliary memory to correspond to the faulty location on the occurrence of the error output and to address the assigned replacement location when the corresponding faulty location of the main store is subsequently addressed by the address register; and gating means connected to receive data outputs from the main memory and the auxiliary memory and operable to select one or other of those data outputs in accordance with whether the location currently addressed by the address register in the main memory is operative or faulty respectively.

One embodiment of the invention will now be described, by way of example, with reference to the accompanying drawing, in which:—

Figure 1 is a schematic block diagram of a memory system, and

Figure 2 is a schematic block diagram of a code conversion arrangement forming part of the memory system.

Data is stored in a plurality of storage locations of a main memory 1 (Figure 1).

Each storage location is addressable and may contain one or more words, or bytes of data. The memory may be engineered in any convenient technology, such as magnetic cores or films, or MOS or other semiconductor elements.

For the purpose of the present description, it will be assumed that the memory is part of a computer system (not shown) and that it is controlled by an address register 2. The register 2 receives from the central processor of the computer system an address word on line 3 and a read or write command on line 4. In the case of a write command, the data on an input line 12 is written into the addressed location of the memory 1, while in the case of a read command the contents of the addressed location are read out on to an output line 5. The data on output line 5 is applied to a gating arrangement 6, which is normally set so as to feed this data to an output highway 7. The highway 7 may be connected to the central processor, for example, so that the selected data may be processed.

It will be assumed that the data in the memory 1 is encoded in one of the known error detecting and correcting codes and the data which is read out on the line 5 is applied to an appropriate data checking device 8. If the device 8 detects an error, it reconstructs the correct data and applies it over line 9 to the gating arrangement 6. It also generates an error signal on line 10 to indicate to the central processor that the data previously read on to the output highway 7 was corrupt and should be replaced by the reconstructed data.

It will be appreciated that the correct data is provided to the central processor, but a longer time is required in comparison with the time normally required for reading out one word, or other unit, of data. Firstly, the correct word has to be reconstructed from the corrupt form and this requires a substantial amount of logical manipulation with many of the usual codes. Secondly, the corrupt word has to be cancelled and replaced by the correct word. This may involve back tracking on one or more instructions of the central processor. If the corruption occurred as a result of a stable fault related to the particular memory location, as opposed to, say a random noise spike, this long reading period will arise each time that the particular location is accessed.

The most likely fault is that one bit position of the storage location will fail. For example, it will not give any output or it will assume a 'stuck at 1' state. Even a relatively simple redundant code will provide for correction of such an error. However, if the fault is in the means by which the storage location is accessed, for example, several of the bit locations may fail. It is likely that correction of such an error will not be possible and the data checking device 8 will signal the central processor over line 11 that it has detected an error which it cannot correct. This signal will stop the program and recovery routines will be initiated either automatically, or by operator intervention, to attempt to circumvent the error condition. Even if this is successful, the faulty location may cause the same situation to arise on the next program which is used.

Furthermore, the occurrence of an error signal on either of lines 10 and 11 may automatically call diagnostic routines which will record details of the error in the main memory, or another store, to assist in the tracing and repair of the fault at some later time. The use of these routines will further increase the time lost by an error. Accordingly, it will be appreciated that leaving faulty locations in use can cause an appreciable reduction in the useful throughput of the computer system, even when the number of faulty locations is quite small.

The effect of the faulty locations is largely eliminated by the use of an auxiliary memory 13 and an associated address code converter 14. The address information from the address register 2 is applied to the address converter 14 over line 15. The address converter is set so that it provides an output address to the memory 13 only if the address which it is receiving is that of a previously detected faulty location in memory 1. In these circumstances, the converter also produces a signal on line 16 which is applied to the gating arrangement 6 to switch it so that the data output from memory 13 on line 17 is fed to the output highway 7, in place of the data output from memory 1, coming from the faulty location. The signal on the line 16 is also applied to the data checking device 8 to suppress its operation, which is not necessary because the correct data will have been read out from the memory 13.

The setting of the address converter and the loading of the correct data into the memory 13 is done in two ways. Firstly, the setting and loading is done periodically, for example, when the computer system is switched on each morning. If any of the memories in the system are volatile, such as semiconductor memories, it will be necessary to initialise them and the loading of the memory 13 can be part of this process.

The setting and loading information is stored in some non-volatile medium such as a disc file, and is read into the memory as part of the initiating process. Thus, the address converter will be set in accordance with all the known faulty locations in the memory 1 and the memory 13 will be loaded with whatever data is currently stored in

those locations. If the data in the faulty locations is changed during the use of the computer system, the corresponding changes will be made in the memory 13, since the input highway 12 is also connected to the memory 13.

The memory system, including memory 1 and memory 13, will appear to be a fault-free system to the central processor, because the output highway 7 will receive the correct data from one or the other memory, so long as no further locations develop faults.

The occurrence of faults during the operation of the computer is dealt with in the following way. The occurrence of a fault causes the data checking device 8 to apply a signal over line 18 to the address converter 14 to set it in accordance with the address which is currently held in the address register 2, which is the address of the location which has shown the fault. The corrected data is applied to the input of the memory 13 from the data checking device 8, so that it will be loaded into a corresponding location of the memory 13.

One form of address code converter which may be used is described in our co-pending application no. 18988/74 (now 1,472,885) and is shown in Figure 2. The address information from the line 15 is applied to an auxiliary address register 20. This auxiliary register is provided primarily to allow the timing of the address converter 14 and memory 13 to be independent of that of the main memory 1.

The more significant half of the auxiliary address register 20 is connected to a random access memory 21 and the less significant half is connected to a similar memory 22. A data output channel of each of the two memories is connected to the address selection inputs of a further random access memory 23. The data inputs of all three memories 21, 22 and 23 are connected via line 24 to the output of a counter 25. The input of the counter 25 is connected to the line 18.

Merely for the sake of illustration, let it be assumed that the memory 1 has sixty-four storage locations which are addressed in decimal code and that the memory 13 has four storage locations. Suppose that the first storage location of the memory 1 which is found to be faulty is location 47. The memories 21, 22 and 23 and the counter 25 are initially set to zero. The auxiliary register 20 will be set to 47, since that is the address in the address register 2. Accordingly, storage location 4 of memory 21 will be addressed by the more significant half of the register 20 and storage location 7 of memory 22, will be addressed by the less significant half of register 20.

The signal on the line 18 which indicates that an error has been detected will step the counter 25 to register one. The signal is also applied to the memories 21 and 22 to test if the addressed storage locations are registering zero. The setting of the counter 25 is transferred to the addressed storage locations, since they are registering zero. Once these storage locations are set, they read out over lines 26 and 27 which are connected to the address selection inputs of the memory 23. Hence, storage location 11 of memory 23 is selected, since each of the storage locations in memories 21 and 22 is generating a one.

The output of the counter 25 is also applied to the memory 23 as a data input, so that one is stored in location 11 of the memory. Once this storage location has been set, it will read out on line 28 to the address selection input of memory 13 to select location 1. Consequently, the corrected data on line 9 from the faulty location 47 of memory 1 is stored in location 1 of memory 13. When storage location 47 of memory 1 is accessed again at some later time, the setting of the register 20 will cause memories 21 and 22 to read out ones. These ones will select storage location 11 of memory 23 to read out a one on line 28. This selects storage location 1 of memory 13, which reads out the appropriate data on line 17. The occurrence of an output from memory 23 generates a signal on line 16. This switches the gate 6 and suppresses operation of the data checking device 8 as already explained.

It is assumed that the memories 21, 22 and 23 have a shorter cycle than the memory 1. Accordingly, the signal on the line 16 will be effective to operate as described when the address converter has already been set. However, when the address converter has not been set previously, the signal on the line 16 will not occur until setting has taken place and this will make it ineffective.

The memories 21, 22 and 23 are preferably self-timing semiconductor memories so arranged that the application of inputs from the register 2 will produce a short period of instability whilst the necessary switching occurs and a stable state is then achieved. However, those skilled in the art will find no difficulty in using other types of memory, the necessary sequence of operations being ensured by employing a suitable set of timing signals, if the memories require synchronisation.

Suppose that the next faulty storage location is location 19. The register 20 will select locations 1 and 9 in memories 21 and 22. These locations are both initially zero and they will be set to two, since the counter 25 will have been stepped on. Accordingly, location 22 will be selected in memory 23 and will be set to two. This will select storage location 2 of the memory 13 to receive the data.

The next faulty location might be 27. The register 20 will select storage locations 2 and 7 in memories 21 and 22, respectively. Storage location 2 will be initially at zero and will be set to three since the counter 25 has been stepped on again. However, storage location 7 is not zero, since it was set to one as a result of the fault on location 47. Accordingly, it is left unchanged so that the previously entered pattern is not destroyed.

As a result, storage location 31 of the memory 23 will be selected and will be set to three to select storage location 3 of memory 13.

It will be clear from the foregoing that the function of the address converter 14 is to assign replacement locations in the auxiliary memory 13 for faulty locations in the main memory 1, and to convert the addresses of the faulty locations, when they appear in address register 2, into the addresses of the corresponding replacement locations.

In a modification of the system described, a single memory may replace the memories 21, 22 and 23. The single memory operates in substantially the same way as the three memories. It is addressed by the register 20 and the storage location corresponding to a faulty address in the store 1 is loaded with the appropriate address of the memory 13. The single memory form of the address converter may be preferable for relatively small sizes of the main store 1, whereas the multi-memory form is better for relatively large sizes. It will be appreciated that the determining factor is the number of input conditions which the memories have to recognise.

In the particular example, a single memory would have to have sixty-four storage locations, corresponding to those of the main memory, with each location being capable of storing any digit from zero to three. In the arrangement shown in Figure 2, the memory 22 requires ten locations (0-9), the memory 21 requires seven locations (0 to 6), and the memory requires sixteen locations (11-14, 21-24, 31-34 and 41-44). Hence, the total number of locations in the multi-memory arrangement is only thirty-three. More than three memories may be employed if the main memory is very big.

The address converter may also be implemented with a content addressed memory device (CAM) in place of the memories 21, 22 and 23. The CAM operates in the conventional manner. The addresses of the faulty locations are stored at the different locations of the CAM and the corresponding addresses in the memory 13 are stored in the associated locations of the CAM. Hence, the address which is entered into the register 20 is compared with all the addresses of faulty locations in the CAM and the corresponding address in memory 13 is read out if a match is found. The use of a CAM is quite satisfactory from a technical point of view, but it may be more expensive to implement.

It will be understood that the memory systems which have been described are of quite general applicability and are not specific to particular functions or structures of the memory. For example, the memory 1 could be part of an indirect addressing system, so that the data which it stores is a directory of addresses in a yet larger storage device, such as a disc file. Equally, the memory 13 could store the addresses at which data is to be found in a larger store, rather than storing the data itself.

When the memory system is operated in a dynamic mode with faulty locations being recorded as they are found, the address decoder and memory 13 hold details of the current fault state of the main store 1. This information is available within the system for the maintenance engineer whenever it may be needed. Furthermore, the fault information is derived from faults occurring under actual operating conditions and it is probably more accurate than the fault information which could be derived from diagnostic routines which must impose artificial conditions to some degree.

It will be appreciated that the memory sizes would normally be very much bigger than has been given in the example. Also, the address selection is more likely to operate on a binary rather than decimal basis. Furthermore, it should be understood that the various data and address lines shown in the drawings are actually multi-bit paths for parallel transmission of data or addresses.

If the memory faults mostly appear at the time of manufacture or shortly thereafter during the testing of the memory system and its associated computer system, the data coding system could be concerned merely with detecting errors, rather than with correcting them. The system as supplied to the computer user would already be set to take account of all the known faults. Even if the necessary routine which is called by detection of an error, is quite complex and time consuming, it will be used very seldom and only once for each fault condition, since the address converter can be up-dated with these additional faulty locations. The simplification of the coding system will save storage space as well as checking equipment. Naturally, the memory system can be added to the existing less sophisticated computer systems which have only error detection features.

WHAT WE CLAIM IS:—

1. A memory system comprising: a main

memory having a plurality of addressable locations each arranged to store an item of data, some of which locations may be faulty; an address register for addressing the main memory to select a storage location; a data checking device responsive to a data item read out from a selected storage location to produce an error output if the data item read out contains an error indicating that the location is faulty; an auxiliary memory having fewer addressable locations than the main memory, each arranged to store an item of data; address code conversion means coupled to the address register and responsive to said error output arranged to assign an address of a replacement location in the auxiliary memory to correspond to the faulty location on the occurrence of the error output and to address the assigned replacement location when the corresponding faulty location of the main store is subsequently addressed by the address register; and gating means connected to receive data outputs from the main memory and the auxiliary memory and operable to select one or other of those data outputs in accordance with whether the location currently addressed by the address register in the main memory is operative or faulty respectively.

2. A memory system according to Claim 1, wherein the data checking device is also arranged to correct data read out of faulty locations in the main memory, the corrected data being stored in the corresponding replacement locations assigned in the auxiliary memory.

3. A memory system according to either preceding claim wherein the address code conversion means includes an arrangement for converting main store addresses into auxiliary store addresses.

4. A memory system substantially as hereinbefore described with reference to Figure 1 of the accompanying drawings.

R. V. P. LOUGHREY,
Chartered Patent Agent,
Agent for the Applicants.

Printed for Her Majesty's Stationery Office by The Tweeddale Press Ltd., Berwick-upon-Tweed, 1977.
Published at the Patent Office, 25 Southampton Buildings, London, WC2A 1AY, from which copies may be obtained.

1491702 COMPLETE SPECIFICATION
1 SHEET

This drawing is a reproduction of the Original on a reduced scale.

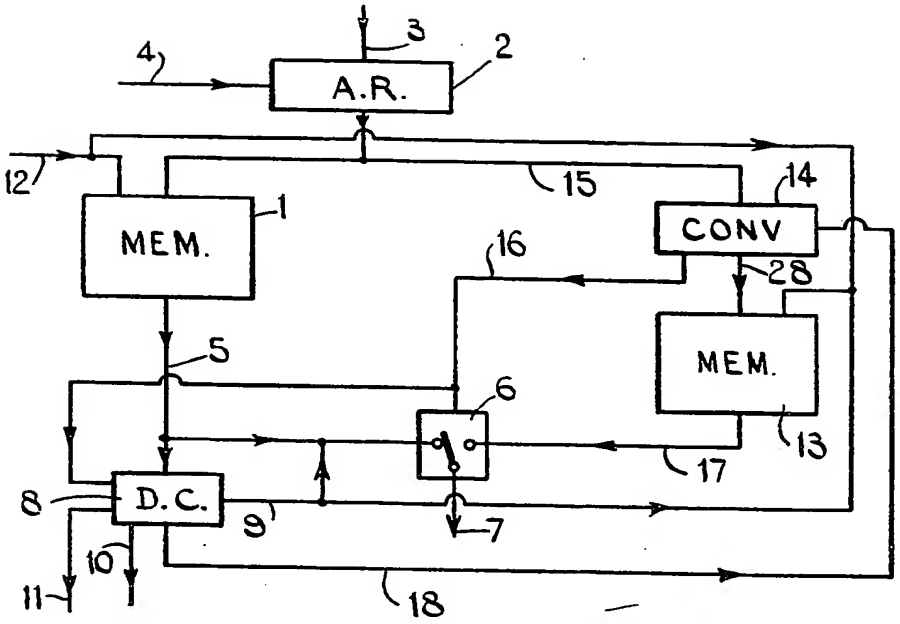

FIG.1.

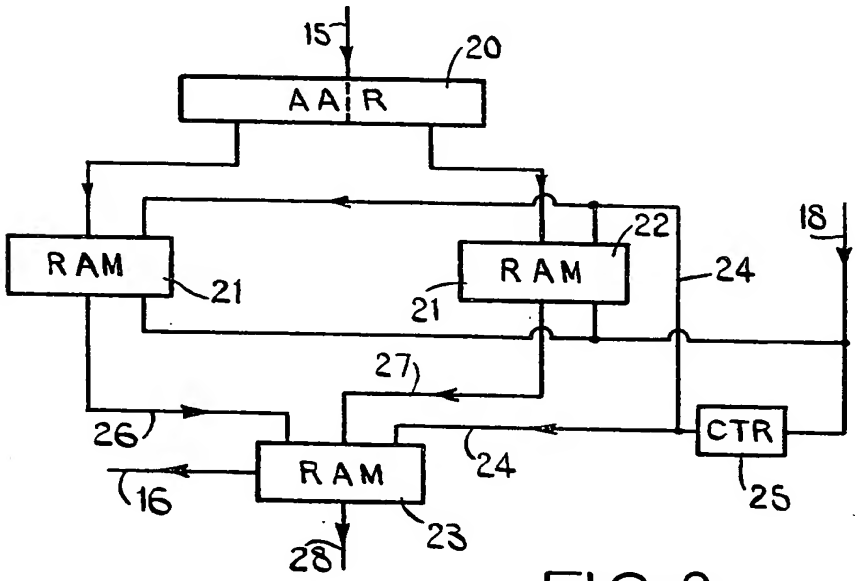

FIG.2.